

What Is Claimed Is:

1. A method comprising:
  - identifying a plurality of candidate summaries related to textual information based at least in part on a document;
  - determining first and second attribute values based at least in part on the candidate summaries; and
  - determining an optimal candidate summary based at least in part on the first and second attribute values.
  
2. The method of claim 1, further comprising:
  - determining a number of storage locations for the plurality of candidate summaries, the number of storage locations based at least in part on a size of the document;
  - combining the plurality of candidate summaries into a plurality of combined candidate summaries;
  - selecting from the plurality of combined candidate summaries a first highest-weighted combined candidate summary and a second highest-weighted combined candidate summary; and
  - comparing the first and second highest-weighted combined candidates, wherein determining the optimal candidate summary is further based at least in part on the comparison of the first and second highest-weighted combined candidate summaries.

3. The method of claim 2 further comprising tokenizing the document.
4. The method of claim 2, wherein the size of the document comprises a number of words disposed in the document.
5. The method of claim 2, wherein the size of the document comprises a number of character strings disposed in the document.
6. The method of claim 2, wherein the number of storage locations is based at least in part on at least one of the size of the document, a desired size of the summary, and a number of query terms.
7. The method of claim 3 further comprising tokenizing a portion of the document sufficient to generate a predetermined number of candidate summaries.
8. The method of claim 1, wherein the first and second attribute values comprise a number of words.
9. The method of claim 1, wherein the first and second attribute values comprise a pixel size.

10. The method of claim 1, wherein the first and second attribute values comprise a string length.
11. The method of claim 2, wherein a number of combined candidates is less than or equal to the number of storage locations.
12. The method of claim 2, wherein each of the storage locations comprises a bucket and wherein the first and second attribute values are mapped onto the bucket.
13. The method of claim 2, wherein the candidate summaries comprise at least one keyword.
14. The method of claim 13, wherein the keyword comprises a first keyword and a second keyword, the optimal candidate summary further based at least in part on a title of the document comprising the first keyword and one of the plurality of candidate summaries comprising the second keyword.
15. The method of claim 13, wherein the first and second highest-weighted combined candidates are further based at least in part on a percentage of keywords included in the plurality of candidate summaries.

16. The method of claim 13, wherein the document comprises a plurality of sections and wherein the first and second highest-weighted combined candidates are further based at least in part on multiple occurrences of the keyword in one of the plurality of sections.

17. The method of claim 13 further comprising adjusting a weighting of the first and second combined candidate summaries based at least in part on a number of the keywords spanned by the plurality of candidate summaries.

18. A method comprising:

- searching a document;
- identifying a keyword disposed in the document;
- identifying a plurality of candidate summaries related to textual information based at least in part on the document;
- determining a number of storage locations for the plurality of candidate summaries, the number of storage locations based at least in part on a size of the document;
- combining the plurality of candidate summaries into a plurality of combined candidate summaries;
- determining first and second attribute values based at least in part on the candidate summaries;

selecting from the plurality of combined candidate summaries a first highest-weighted combined candidate summary and a second highest-weighted combined candidate summary;

determining an optimal candidate summary based at least in part on the first and second attribute values; and

comparing the first and second highest-weighted combined candidate summaries, wherein determining the optimal candidate summary is further based at least in part on the comparison of the first and second highest-weighted combined candidate summaries.

19. A computer readable medium comprising instructions, that, when executed, cause an application to:

identify a plurality of candidate summaries related to textual information based at least in part on a document;

determine first and second attribute values based at least in part on the candidate summaries; and

determine an optimal candidate summary based at least in part on the first and second attribute values.

20. The computer readable medium of claim 19, further comprising instructions, that, when executed, cause an application to:

determine a number of storage locations for the plurality of candidate summaries, the number of storage locations based at least in part on a size of the document;

combine the plurality of candidate summaries into a plurality of combined candidate summaries;

select from the plurality of combined candidate summaries a first highest-weighted combined candidate summary and a second highest-weighted combined candidate summary; and

compare the first and second highest-weighted combined candidate summaries, wherein determining the optimal candidate summary is further based at least in part on the comparison of the first and second highest-weighted combined candidate summaries.

21. The computer readable medium of claim 19, further comprising instructions, that, when executed, cause an application to tokenize the document.

22. The computer readable medium of claim 19, wherein the size of the document comprises a number of words disposed in the document.

23. The computer readable medium of claim 19, wherein the size of the document comprises a number of character strings disposed in the document.

24. The computer readable medium of claim 20, wherein the number of storage locations is based at least in part on at least one of the size of the document, a desired size of the summary, and a number of query terms.
25. The computer readable medium of claim 19, wherein the first and second attribute values comprise a number of words.
26. The computer readable medium of claim 19, wherein the first and second attribute values comprise a pixel size.
27. The computer readable medium of claim 19, wherein the candidate summaries comprise a string length.
28. The computer readable medium of claim 20, wherein a number of combined candidate summaries is less than or equal to the number of storage locations.
29. The computer readable medium of claim 20, wherein each of the storage locations comprises a bucket and wherein the first and second attribute values are mapped onto the bucket.
30. The computer readable medium of claim 20, wherein the candidate summaries comprise at least one keyword.

31. The computer readable medium of claim 30, wherein the keyword comprises a first keyword and a second keyword, the optimal candidate summary further based at least in part on a title of the document comprising the first keyword and one of the plurality of candidate summaries comprising the second keyword.

32. The computer readable medium of claim 30, wherein the first and second highest-weighted combined candidates are further based at least in part on a percentage of the keywords included in the plurality of candidate summaries.

33. The computer readable medium of claim 30, wherein the document comprises a plurality of sections and wherein the first and second highest-weighted combined candidates are further based at least in part on multiple occurrences of the keyword in one of the plurality of sections.

34. The computer readable medium of claim 21 further comprising instructions, that, when executed, cause an application to tokenize a portion of the document sufficient to generate a predetermined number of candidate summaries.

35. The computer readable medium of claim 30 further comprising instructions, that when executed cause an application to adjust a weighting of the first and second combined candidate summaries based at least in part on a number of the keywords spanned by the plurality of candidate summaries.